

Causal discovery of drivers of surface ozone variability in Antarctica using a deep learning algorithm

Pankaj Kumar

Indian Institute of Technology Kharagpur

Dr. Jayanarayanan Kuttippurath (✉ jayan@coral.iitkgp.ac.in)

Indian Institute of Technology Kharagpur

Dr. Adway Mitra

Indian Institute of Technology Kharagpur

Research Article

Keywords: Causal Discovery, ozone, tropospheric ozone, pollution, teleconnection, antarctica, causal inference, causal graph

Posted Date: August 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-826824/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Environmental Science: Processes & Impacts on January 1st, 2022. See the published version at <https://doi.org/10.1039/D1EM00383F>.

1 **Causal discovery of drivers of surface ozone variability in**
2 **Antarctica using a deep learning algorithm**

3 P. Kumar¹, J. Kuttippurath^{1,✉}, and A. Mitra²

4

5 ¹ CORAL, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India

6 ² Centre of Excellence in Artificial Intelligence, Indian Institute of Technology Kharagpur,
7 Kharagpur, West Bengal 721302, India

8 ✉ Correspondence: J. Kuttippurath <jayan@coral.iitkgp.ac.in>

9

10 **Abstract**

11 The discovery of causal structures behind a phenomenon under investigation has been at the
12 heart of scientific inquiry since the beginning. Randomized control trials, the gold standard for
13 causal analysis, may not always be feasible, such as in the domain of climate sciences. In the
14 absence of interventional data, we are forced to depend only on observational data. This study
15 demonstrates the application of one such causal discovery algorithm using a neural network for
16 identifying the drivers of surface ozone variability in Antarctica. The analyses reveal the
17 overarching influence of the stratosphere on the surface ozone variability in Antarctica,
18 buttressed by the southern annular mode and tropospheric wave forcing in mid-latitudes. We find
19 no significant and robust evidence for the influence of tropical teleconnection on the ground-
20 level ozone in Antarctica. As the field of atmospheric science is now replete with a massive
21 stock of observational data, both satellite and ground-based, this tool for automated causal
22 structure discovery might prove to be invaluable for scientific investigation and flawless decision
23 making.

24 **Introduction**

25 Ubiquitous throughout the troposphere and stratosphere, ozone plays a significant role in
26 atmospheric radiative forcing, atmospheric chemistry, and air quality. Considered an
27 atmospheric cleanser, Ozone in the stratosphere (90% of total amount) saves life on Earth by
28 filtering harmful UV radiation. Stratospheric ozone throughout the globe has been on a
29 downward trend, as indicated by the analysis of both satellite and ground-based measurements of
30 total column ozone (TCO) due to a steady increase in anthropogenic emissions of the reactive
31 chlorofluorocarbons (CFCs).¹ However, the stratospheric ozone hole is on a recovery path in
32 response to the Montreal Protocol and its subsequent amendments.²⁻⁴ Nonetheless, the precise
33 causes of the observed changes in stratospheric ozone are complicated to isolate. They remain
34 uncertain due to the inability of existing chemistry-climate models (CCMs) to reproduce the
35 observations.

36 In contrast, tropospheric ozone is a prominent air pollutant and greenhouse gas despite being
37 only 10% of the total column amount. The accurate assessment of tropospheric ozone trends is
38 far more challenging due to the complicated interplay of many simultaneous processes with
39 significant temporal and spatial variations. Ground-level ozone concentration at a given location
40 is affected by photochemical reactions, atmospheric transport, atmospheric diffusion,
41 topography, and emission sources of the primary pollutants [such as nitrogen oxides (NO_x) and
42 non-methane volatile organic compounds (NMVOCs)]. The inter-annual variability of ozone
43 concentration is governed by changes in emission of photochemical precursors, various
44 favorable/unfavorable weather conditions, or a combined effect of all these. Favorable
45 meteorological conditions for high ozone episodes include high temperature, intense sunlight,
46 and light wind.⁵ On the other hand, dry deposition, dissolution into the seawater, and photolysis
47 reactions involving nitrogen oxides (NO_x) are the most prominent sinks of tropospheric ozone.

48 A variety of methods has been applied to date for the analysis of ground-level ozone that ranges
49 from simple statistical models like multiple linear regression (MLR) to sophisticated chemistry-
50 climate models (CCMs) such as GEOS-Chem.⁶ However, these models face difficulty in dealing
51 with complicated cause-effect relationships among meteorology and air pollutants. Multiple
52 regression models are limited in their interpretability as these are based on cross-correlation,
53 which might be highly biased due to autocorrelation effects or spurious correlations arising from
54 an unaccounted third process or a common driver. Apart from these, they lack insights into the
55 directionality of relationships. Therefore, CCMs are used to investigate the impact of changes in
56 emissions and meteorology using controlled perturbation of the system, allowing interpretation
57 of simulation results as causal effects forced by the interventions. Nonetheless, the ability of
58 CCMs in resolving essential processes such as land–biosphere interactions, stratosphere-
59 troposphere transport (STT), and detailed atmospheric dynamics remain questionable, restricting
60 their interpretability and conclusions.⁷

61 Causality is a fundamental scientific notion and is indispensable for accurate forecasting,
62 flawless explanation, and decision making. Discovering causal relations from observational data
63 has drawn much attention recently as the traditional way of causal analysis using interventions or
64 randomized control trials might be impractical, infeasible, or outright unethical. For example,
65 causal discovery methods relying solely on observational data have been used recently to study

66 ocean-atmosphere interactions,⁸ the Walker circulation,⁹ and the mid-latitude winter circulation
67 in the northern hemisphere.⁷ With methodologies based on conditional independence tests,
68 heuristic scoring, or deep learning, we can identify causal linkages in observational data using
69 the premise that causes temporally precede their effects in time series. In this paper, we use one
70 such causal model based on a deep neural network to discover the potential drivers of surface
71 ozone variability over Antarctica. This method overcomes the pitfalls of common statistical
72 approaches, i.e., spurious correlations arising from the presence of common drivers,
73 autocorrelation, or indirect effects using a carefully designed causal discovery algorithm.

74 **Results**

75 Fig. 1 shows the daily time series of surface ozone measurements at all five stations considered
76 in this study, i.e., Arrival Heights, Marambio, Neumayer, South Pole, and Syowa. As shown in
77 the figure, surface ozone has a marked seasonal cycle with the highest concentration during
78 winter (June-July-August; JJA) and the lowest during summer (December-January-February;
79 DJF); consistent with the remoteness of the Antarctic continent. However, there are a few
80 noticeable differences in surface ozone seasonal cycle among various stations. There is a clear
81 secondary peak (up to 50 ppbv) during spring (September-October-November; SON) at South
82 Pole station recurring every year and have concentrations equivalent or higher than those during
83 the primary peak (JJA). In contrast, secondary peaks at all other stations are sporadic and rarely
84 exceed those during the primary peak. The occurrence of the secondary peak in Antarctica has
85 been attributed to enhancement episodes due to NO_x emission from snowpack¹³ and photolysis
86 of remote PAN formed above continental source regions upon descent within the Antarctic
87 region.³⁰ Notwithstanding, these peaks might also result from the transport of photochemically
88 produced ozone in the planetary boundary layer (PBL) over the Antarctic plateau to other parts
89 of Antarctica due to katabatic flow prevalent apart from the direct transport of air mass from
90 UTLS enriched in ozone.

91 We identify the ozone enhancement events [OEE] at all stations included in this study using the
92 methodology adopted by Cristofanelli et al. 2018.³¹ OEEs identified at all stations in Antarctica
93 are shown in Fig. 1 in magenta color. To identify the OEE, we first fit an annual sinusoidal curve
94 (green curve in the top panel of Fig. 1) to the daily surface ozone dataset, followed by estimation

95 of gaussian PDF of residual (grey curve in the top panel of Fig. 1) from sinusoidal fit. We fit
96 another Gaussian distribution to all points lying beyond one σ of the last PDF. The intersection
97 of these two PDFs (vertical dash line in the bottom panel of Fig. 1) shows the threshold value for
98 an OEE event. Analyses suggest that OEEs primarily occur during late spring and early summer
99 (November-December-January) at all stations with the highest frequency during December
100 month at the South Pole (up to 80%) and November month at coastal stations (60%, 70%, and
101 65% at Arrival Heights, Neumayer, and Syowa respectively) as shown in Fig. S1. On an annual
102 basis, OEE frequency ranges from 2% at Arrival Heights to 25% at the South Pole. OEE
103 frequency has been increasing at South Pole, Arrival Heights, and Neumayer since 1990,
104 whereas it has been decreasing at Syowa.

105 **Potential drivers of ozone variability**

106 Several studies have attributed the changes in surface ozone to short-range and long-range air
107 mass transport and stratosphere-troposphere exchange (STE).^{32–35} Atmospheric waves
108 propagating from the troposphere to the stratosphere control the variation in STE, which is in
109 turn modulated by the meridional circulation driven by momentum imparted by the waves. Fig. 2
110 (top panel) shows the time series of inter-annual variation in STT at all five stations in
111 Antarctica. As shown, STT contributes up to ~15% of the air mass transport at surface level in
112 Antarctica, with a considerable variation across different stations. STT is the highest during
113 March (Autumn) at all stations. There is enhanced STT contribution during the summer and
114 autumn seasons as compared to winter and spring. The South Pole receives the highest
115 proportion of air mass transport from the stratosphere as it is situated on the Antarctic Plateau,
116 and Marambio receives the least. Arrival Heights receives the highest amount of STT among
117 coastal stations due to the influence of katabatic winds blowing from the high-altitude plateau
118 region, consistent with previous studies.³⁶ As shown in Fig. 2, the contribution of STT increased
119 during 1980–2000 and is decreasing since then.

120 We now investigate the influence of descending air mass transport from the UTLS on OEE
121 variability in Antarctica. Fig. 2 (b and c) shows the interannual variation in the frequency of
122 OEEs during the spring and summer seasons (the seasons with the highest number of OEEs in a
123 year). The frequency of OEEs during these two seasons had an increasing trend during the
124 1990s. However, this increasing trend has been disrupted during recent years and has turned to

125 decreasing trend during summer. Fig. 2 (d and e) shows the interannual variation in the
126 frequency of the fraction of 15 days back-trajectories coming from the upper troposphere
127 corresponding to OEEs. It shows a high contribution (40–100% of OEEs) of air mass coming
128 from the upper troposphere (with pressure less than 500 hPa). Downward transport to the lower
129 troposphere from UTLS is driven by residual mean flow due to Rossby wave forcing on the
130 poleward side of the jet stream, facilitating a secondary circulation from the vortex edge to the
131 inner vortex observed over the South Pole and Arrival Heights.^{37,38}

132 A close analysis of OEEs during summer suggests the evident influence of southern annular
133 mode (SAM) on the occurrence of OEEs, as shown in Fig. 2 (c). The frequency of OEEs across
134 all stations in Antarctica seems to follow the variation in the SAM index. After 1997, there is a
135 clear increasing trend in OEEs during the period in which the SAM index has been decreasing
136 and vice-versa. We estimate the R^2 between SAM index and OEEs during 1997–2015 summer
137 across all stations. OEEs at Marambio, South Pole, and Arrival Heights have very strong
138 influence from SAM. It explains ~93%, ~68%, and ~61% of OEEs variance at these stations,
139 respectively. Similarly, it explains ~27% and ~20% of OEEs variance at Neumayer and Syowa,
140 respectively. SAM drives changes in mid-latitude jet, which modulates the lower tropospheric
141 circulation, causing variations in surface ozone concentration. As the variation in SAM and
142 southern mid-latitude jet position has been previously shown to be driven by stratospheric ozone
143 hole in Antarctica,⁴ it demonstrates the clear impact of the stratospheric ozone hole and
144 stratosphere-troposphere coupling on summertime OEE variation. This pattern in OEE variation
145 is consistent with previous studies wherein stratospheric cooling associated with ozone loss in
146 the stratosphere and increasing greenhouse gases were found to shift the tropopause altitude and
147 transport ozone to the troposphere with a lag of about a month.^{39,40} As the polar vortex
148 strengthened due to increased ozone loss in the stratosphere and global warming, it caused
149 enhanced transport from UTLS to the lower troposphere resulting in an increase in the frequency
150 of occurrence of OEEs.

151 Furthermore, atmospheric transport over the Antarctic continent is controlled by several
152 synoptic-scale pressure patterns. For example, Amundsen–Bellingshausen Seas low (ABSL) has
153 substantial control over the tropospheric circulation in the Ross Sea region.³⁷ Similarly, lows
154 over the drake passage, Amundsen, Bellingshausen, Weddell Seas, and the ridge over the

155 peninsular region drive the air transport over the Antarctic Peninsula. These are influenced by
156 SAM and El Niño and southern oscillation (ENSO) in turn.⁴¹ Moreover, meridional circulation is
157 also influenced by ENSO and Quasi-biennial oscillation (QBO). It can make substantial
158 contributions to the transport of ozone from the stratosphere to the troposphere. Therefore, we
159 include different factors that are known to control ozone variability in the troposphere,^{42,43} i.e.,
160 solar flux (SF) at 10.7 cm wavelength, heat flux (HF) at 200 hPa (averaged over 45°–75°S)
161 accounting for the changes in meridional circulation, potential vorticity (PV) at 200 hPa
162 accounting for variation in the strength of polar vortex and STT, and aerosol optical depth
163 (AOD) at 550 nm (averaged over 45°–75°S) which represents aerosol loading in the atmosphere
164 and volcanic eruptions in addition to the SAM index, multivariate ENSO index (MEI) and the
165 equatorial winds at 30 and 50 hPa which represent QBO.

166 Table S2 shows the maximum cross-correlation of selected proxies with surface ozone at various
167 stations. As is clear from the table, the selected proxies are highly correlated with surface ozone
168 and thus should be able to explain the surface ozone variability. The MLR fit with these proxies
169 at all stations is performed to find out the usefulness of selected proxies for explaining the
170 surface ozone variability and are shown in Fig. 3 (without seasonal harmonics) and Fig. S2 (with
171 seasonal harmonics). MLR suggests that selected variables can explain the variations in the
172 surface ozone reasonably well (with R_{adj}^2 ranging from 0.9 in the case of the South Pole to 0.96
173 for Neumayer and Syowa).

174 The LASSO fit for surface ozone at Neumayer is shown in Fig. S5, and they show a good fit
175 with R^2 of 0.58 at Neumayer. The estimated regression coefficients for all proxies suggest that
176 PV, HF, AAO, SF, and MEI are the only important factors for explaining the surface ozone
177 variability, with PV being the most crucial factor accounting for the majority of variations. While
178 LASSO regression allows discovering active variables, it does not deal well with the strong
179 inter-dependencies due to the spatio-temporal nature of the variables. Multiple regression models
180 based on cross correlation (both MLR and LASSO) can be strongly biased by autocorrelation
181 effects, indirect linkages via a third process, or a shared driver leading to noncausal, false
182 correlations that restrict their interpretability. Furthermore, it does not provide any information
183 on the direction of the link, making it inadequate for studying causal effects.

184 **Causal Discovery of the drivers**

185 Having identified the inadequacies of MLR and LASSO, we now apply our proposed TCDF
186 framework to identify the causal drivers of surface variability and compare it with the same done
187 using causal effect network (CEN) analysis using PCMCI. Herein, we examine whether the
188 surface ozone variability is driven by stratospheric variability and teleconnections with various
189 climate modes such as ENSO, QBO, and AAO. The usage of TCDF for surface ozone at
190 Neumayer station is illustrated in Fig. S6. In brief, TCDF tries to identify the causal drivers of
191 surface ozone variability by regressing it with the included proxies (shown in Fig. S6 left
192 column) using attention based 1-D Temporal CNN. The identified potential drivers are then
193 subject to further validation by using PIVM to identify the true causes. Then a causal network
194 graph is generated, as shown in Fig. S6 right panel.

195 Discovered causal relationships for all four stations are shown in Fig. 4. Our analyses suggest
196 that surface ozone is influenced by changes in stratosphere and Brewer-Dobson circulation
197 through changes in PV at 200 hPa. Although, no indication of dependence of surface ozone on
198 tropical teleconnections is found. SAM has significant control over the surface ozone at all four
199 stations. Similarly, we do not find any influence of the solar cycle on surface ozone variability.
200 To confirm the discovery made by TCDF, we fit another MLR on surface ozone using the
201 discovered causal drivers only, and the MLR fit for the same is shown in Fig. S3. It clearly
202 shows that we can achieve the same fit (shown by R^2) as done before without using other non-
203 causal variables, suggesting the irrelevance of using non-causal variables in explaining ozone
204 variability. We tried another MLR analysis (Figure S4) using non-casual variables, and it shows
205 a considerable reduction in R^2 .

206 Having discovered the causal parents of surface ozone, we now estimate their causal effect on
207 surface ozone variability by calculating their average causal effect (ACE) on surface ozone at all
208 stations. ACE consists of both direct and indirect effects of causes on surface ozone. ACE has
209 been estimated using gAIPW in the framework of the potential outcomes, and are shown in
210 Table 1. The ACE can be interpreted as the changes in surface ozone with the corresponding
211 change of 1 standard deviation (sd) in the geophysical driver under investigation. PV (0.58–0.93)
212 and HF (0.12–0.24) have a positive relationship with surface ozone at all stations included in this

213 study. In contrast, AAO (0.13–0.32) and AOD (2.51–3.75) have a negative impact on surface
214 ozone in Antarctica.

215 To validate the causal relationships discovered by TCDF, we perform another causal discovery
216 using causal effect network (CEN) analysis with a different algorithm. CEN uses PCMCI to
217 detect causal relationships by forming networks between variables under consideration at
218 different time lags. It evaluates the partial correlation between different combinations of
219 variables iteratively conditioning for other variables and their parents. Those found significant
220 even after accounting for other variables are retained. The identified causal links (at 95%
221 confidence interval) for surface ozone variability at different stations are shown in Fig. 5. Here, a
222 significant positive causal relation is shown with a red arrow, whereas the blue arrows show a
223 negative correlation. Links without the arrows show a strong correlation, but the causal
224 relationship cannot be determined with the given information because of coarser sampling. In
225 contrast to TCDF, the CEN analyses reveal that all tested geophysical drivers are causally
226 correlated with surface ozone at 95% confidence interval either directly or indirectly but at
227 different time lags. The discovered relationships are consistent across the stations except for
228 solar flux, which has a significant negative causal relationship with surface ozone only at
229 Neumayer.

230 **Discussion**

231 This study uses PCMCI and a deep learning-based causal discovery framework, TCDF, to
232 diagnose the causal drivers of surface ozone variability in Antarctica. Causal effects of
233 discovered drivers using TCDF have been estimated using a doubly robust estimator based on
234 the potential outcome framework. Our analyses indicate that the surface ozone variability at all
235 four Antarctic stations is driven primarily by stratospheric variability (PV and HF at 200 hPa).
236 The surface ozone is bereaved of teleconnections with tropical climate variabilities such as
237 ENSO and QBO, although a strong influence with AAO is found. Similarly, the solar cycle does
238 not seem to have any influence over the surface ozone variabilities. The relationship of PV with
239 surface ozone is confounded by HF, or PV mediates the influence of HF on surface ozone.
240 Likewise, the relationship of AOD with ground-level ozone is confounded by AAO. The causal

241 links discovered by TCDF are more sparse than PCMCI, which discovered many more links and
242 is quite vague about the impact of climate modes on surface ozone.

243 The stratospheric control of surface ozone variabilities in Antarctica mediated by the heat flux in
244 the UTLS region is consistent with our understanding of the influence of stratospheric variability
245 inside polar vortex over the surface climate, which in turn affect the surface ozone variations.⁴⁴
246 For example, Boljka et al.⁴⁵ found that a weakened tropospheric zonal flow tends to be preceded
247 by stratospheric warmings forced by tropospheric wave sources (both near the tropopause and
248 the surface), which in turn might drive the surface ozone variability. Similarly, Wang et
249 al. 2021³⁷ reported a robust association between Antarctic sea ice and stratospheric polar vortex
250 variability in both observations and model simulations, mediated by Amundsen Sea low and
251 surface winds changes. Among the climate modes considered, only AAO seems to have a
252 significant impact on surface ozone. Our analyses reveal the coupling between ENSO, AAO, and
253 QBO, consistent with previous studies, where a similar interaction between ENSO, AAO, and
254 QBO is observed. For example, Pohl et al.⁴⁶ and Carvalho et al.⁴⁷ suggest that El Niño
255 corresponds to negative AAO phase, as observed in both TCDF and CEN analyses at the South
256 Pole, whereas Taguchi et al.⁴⁸ present the evidence of co-variation of QBO and ENSO.

257 Our results are consistent with Lu et al.⁴⁹, wherein they analysed the tropospheric ozone
258 variability in the southern hemisphere (SH) using GEOS-Chem. They reported the poleward
259 expansion of the SH Hadley circulation (SHHC) to be responsible for the tropospheric ozone
260 increase. As shown by various studies, poleward expansion of SHHC is synonymous with the
261 increasing trend in AAO as the AAO confounds the latitudinal position, width, and strength of
262 mid-latitude jet.⁴

263 While the causal graphs generated by TCDF are sparse and interpretable, they are not entirely
264 devoid of modeling errors and require domain knowledge to identify discrepancies. For example,
265 the interactions among ENSO, AAO, and QBO vary across different stations as analysed using
266 TCDF. In contrast, they are same across all stations as analysed using PCMCI. These caveats
267 might be handled better using a more advanced causal discovery algorithm that leverages shared
268 dynamics across different causal graphs and robustly deals with hidden confounders to discover
269 causal links from time-series data like Amortized Causal Discovery.⁵⁰

270 In summary, we perform the causal analysis of surface ozone variability in Antarctica using a
271 state-of-the-art causal discovery framework based on a deep temporal convolutional network.
272 This framework avoids the drawbacks of common multivariate regression methods and generates
273 a causal graph that is sparse and interpretable. The generated causal graphs were found to be
274 consistent with the existing knowledge. With exponential growth in the amount of observational
275 data from both satellite and ground-based measurements, causal discovery methods might
276 provide novel insights across various domains of atmospheric and climate sciences, which can
277 aid knowledge discovery and guide robust policymaking.

278 **Methods**

279 **Data**

280 In this study, ground-based surface ozone measurements from 5 Antarctic stations, namely
281 Arrival Heights, Marambio, Neumayer, South Pole, and Syowa, are used (see Table S1 for
282 details). While surface ozone measurements at Arrival Heights and Syowa start from 1997, those
283 at the South Pole and Neumayer start much earlier. However, there was an instrument change at
284 Neumayer during 1992, which produced a marked difference in surface ozone measurements.
285 Therefore, we have taken 1993 as the starting year for this study. Surface ozone measurements at
286 different stations have data gaps, i.e., the South Pole measurements have gaps during
287 September–December 2016 and August–December 2017. Similarly, Arrival Heights station data
288 has missing data during October–December 2016 and December 2017. Henceforth, we have used
289 monthly surface ozone during 1997-2015 for causal discovery to circumvent the data gaps. As
290 the measurements at Marambio started late, we do not consider Marambio measurements for
291 causal discovery and inference.

292 **Estimation of stratosphere-troposphere transport**

293 We use a lagrangian transport model HySPLIT using meteorological data from National Center
294 for Environmental Prediction (NCEP) (2.5° latitude-longitude grids) and Global Data
295 Assimilation System (GDAS) (1° latitude-longitude grids) to generate 15 days backward
296 trajectories on a daily basis at 500m above the ground level [agl]. Both meteorological datasets

297 have been used widely in several studies concerned with the airmass transport in Antarctica.
298 They can capture the meteorological variability in the Antarctic region reasonably well.¹⁰⁻¹³
299 Generated backward trajectories have been discretized to 1° x 1° latitude-longitude grids. After
300 that, we use NCEP tropopause data to identify the trajectories coming from the stratosphere. Any
301 trajectory with endpoints with corresponding pressure lower than the associated grid tropopause
302 pressure is marked as being influenced by stratospheric transport and is counted over each month
303 to estimate the monthly frequency of stratosphere-troposphere transport (STT).

304 **Causal Discovery**

305 The goal of causal discovery is to uncover causal relationships using observational data. Before
306 finding causal relationships between distinct combinations of drivers at different time delays,
307 causal discovery methods must overcome numerous hurdles provided by the causative process or
308 the sampling process generating observational data. Because the cause typically occurs before
309 the effect, utilising the concept of time aids in the determination of the directionality of a causal
310 link. Causal links are the relationships found to be significant even after accounting for the
311 influences of other drivers (observed or hidden) or auto-correlations.

312 **Granger Causality**

313 Testing time-lagged causal connections in the framework of Granger causality (GC) is a popular
314 method to causal discovery. A widely used approach to identify the drivers of atmospheric ozone
315 variability utilizing a linear GC framework is to use an autoregressive regression model.¹⁴⁻¹⁷ The
316 generalized regression models (e.g. MLR) assume a linear relationship between the quantity
317 under investigation and selected exogenous quantities (predictors). Mathematically, one such
318 MLR model for surface ozone variability in Antarctica can be represented as follows:

$$319 \quad Y_t = c + xt + \sum_{n=1}^4 (a_n \cos(n\omega t) + b_n \sin(n\omega t)) + \sum_i q_i F_i + \epsilon_t \quad (1)$$

320 where $\omega = 2\pi/12$ for surface ozone observations sampled monthly.

321 This MLR model has the following components: a constant mean level (represented by
322 regression coefficient c) and a linear trend (x), seasonal effects (a_n and b_n), components to
323 describe the influence of external forcings (F), and noise (ϵ) with autoregressive correlation. It

324 includes an autoregressive noise factor to account for irregular cycles, long-range dependencies,
325 and the impacts of various driving mechanisms that a model overlooks. The parameters of the
326 model are estimated by performing least-square minimization.¹⁸

327 If the predictor variables are highly correlated, multicollinearity can become a problem during
328 MLR fitting, causing the coefficient estimates of the model to be unreliable and have high
329 variance. The least absolute shrinkage and selection operator (LASSO) regression overcomes
330 this issue by adding a penalty term to the traditional MLR model objective. It produces a sparse
331 model with only a subset of the input predictor variables, enhancing the prediction accuracy and
332 interpretability of the resulting statistical model.¹⁹ Mathematically, the minimization objective of
333 a LASSO model, i.e., penalized residual sum of squares (PRSS), can be represented as follows:

$$334 \quad PRSS(\beta) = MLR_{obj} + \lambda \sum_i |\beta_i| \quad (2)$$

335 where β_i represents regression coefficients, i.e., a_n , b_n , and q_i in MLR and λ represents the lasso
336 penalty ($\lambda \geq 0$).

337 **Pearl Causality**

338 Linear Granger causal methods like lasso regression can only identify causal relationships with
339 observed data. In practise, however, not all key parameters may be observed, and GC is unable to
340 appropriately deal with unmeasured time series, including hidden confounders. Low detection
341 power of linear GCs can be accentuated using a constraint-based approach that uses a series of
342 independence tests to identify causal links, including hidden confounders. PCMCI is one such
343 method, the most widely used causal discovery algorithm in climate sciences.^{7,20–23} The PCMCI
344 algorithm (working within the premises of Pearl causality (PC)) consists of two steps:

- 345 1. Identification of parents of each driver using Peter Spirtes and Clark Glymour (PC)-
346 algorithm based condition selection which performs iterative conditional independence test
347 by calculating the partial correlation between two time series conditioning on other
348 available time-series at different lags.

349 2. Evaluation of the significance of causal links by determining p-values using momentum
350 conditional independence (MCI) test followed by calculating the strength of causal links
351 using MLR.

352 PCMCI algorithm has two free parameters, which have to be chosen by the user: maximum time
353 delay (τ_{\max}) and significance threshold (α). Further details about the PCMCI can be found in
354 Runge et al. 2015.²⁴

355 **Temporal Causal Discovery Framework (TCDF)**

356 Despite the high detection power of PCMCI, it cannot detect the contemporaneous links and is
357 suited for assessing linear relationships only. In addition, it requires constraining pre-conditions
358 rarely satisfied in climate sciences like stationarity. However, several novel causal discovery
359 methods have been proposed recently circumventing these requirements. For example, Neural
360 Additive Vector Autoregression (NAVAR) extends the popular Vector Autoregression (VAR) of
361 GC framework to nonlinear additive relationships modeling²⁵ and uses a deep neural network
362 (DNN) to do the same. Tank et al. 2021²⁶ extended the LASSO regression in order to perform
363 non-linear causal discovery with the help of a neural network based on Multi-layer perceptron
364 (MLP) and Long-short term memory (LSTM) using convex group-lasso penalties. However,
365 these methods based on GC and PC framework struggle with the presence of hidden
366 confounders.

367 Nauta et al. 2019²⁷ proposed a novel causal discovery framework called Temporal Causal
368 Discovery Framework (TCDF), which uses attention-based temporal convolutional networks
369 (TCNs) for identifying non-causal links, including the hidden confounders. TCDF performs the
370 causal discovery in three stages:

371 1. **Identification of potential causes:** TCDF uses multiple one-dimensional convolutional
372 networks (CNNs) called Attention-based Dilated Depth-wise Separable Temporal
373 Convolutional Networks (AD-DSTCNs) to identify the potential causes. Herein, multiple
374 layers of TCN control the allowable temporal lags among input time series.

375 2. **Causal Validation and Delay Discovery:** Discovered potential causes in the first stage are
376 validated using Permutation Importance Validation Method (PIVM), i.e., original time

377 series is intervened by random permutation, and the resulting loss is compared with the loss
378 estimated using original time series to isolate the true causes.

379 3. **Construction of causal graph:** Finally, TCDF interprets the internal parameters of TCNs
380 and summarises the discovered causal linkages between input time series and relevant time
381 delays by constructing a causal graph.

382 TCDF has a few hyperparameters such as the number of epochs, number of hidden layers, kernel
383 size, dilation coefficient, loss function, significance level for intervention loss, and learning rate.

384 Here, we perform the causal discovery using TCDF due to its simplicity and the ability to deal
385 with the hidden confounders. Our causal discovery method utilizing the TCDF framework is
386 different from that of Nauta et al. 2019²⁷ in the sense that the algorithm is not constrained to look
387 for the cutoff attention score (τ_j) in the first half during the attention interpretation stage as it
388 would restrict the number of potential causes to just half of all time series included in the study.
389 Discovered causal graphs from TCDF are compared with the same derived using PCMCI to
390 ascertain its robustness. We have also performed the MLR and lasso regression to identify the
391 inadequacies of these traditional statistical techniques for causal discovery.

392 We include various proxies (F_i) representing exogenous processes that drive the changes in
393 surface ozone in Antarctica. As surface ozone has substantial seasonal variability, we include
394 regression parameters expressed by a cosine and sine harmonic expansion utilising four
395 harmonics, i.e., 12 months ($n = 1$), 6 months ($n = 2$), 4 months ($n = 3$), and 3 months ($n = 4$)
396 in eq. 1. All data are rescaled to $[-0.5, 0.5]$ before use for modeling. Since we are interested in
397 determining the drivers of surface ozone variability, we also estimate the adjusted coefficients of
398 determination (R_{adj}^2) for MLR as it gives a measure of improvement in model fit when a
399 parameter is added to the model.¹⁷

400 We test for the stationarity of our datasets using Kwiatkowski–Phillips–Schmidt–Shin (KPSS)
401 test and Augmented Dickey-Fuller (ADF) unit root test before performing causal discovery using
402 PCMCI and stationarise the dataset by first order differencing as required. We take six months as
403 τ_{max} to account for tropical teleconnections to the polar region, and α is taken as 0.05.

404 **Estimation of causal effects**

405 We estimate the average causal effect (ACE) following the framework of potential outcomes.
406 The causal effect is defined as the difference between two potential outcomes. Here, the first
407 potential outcome concerns the treatment group and the other with the intervention or control
408 group.⁹ Analytically, ACE is defined as:

$$409 \quad ACE(x_1, x_2) = E[Y|do(X = x_1)] - E[Y|do(X = x_2)] \quad (3)$$

410 where $do(X = x)$ represents an intervention that sets X to x .

411 As both potential outcomes cannot be observed simultaneously, the strong ignorability
412 assumption is required to identify causal effects. Since the causal graph must be known
413 (existence and absence of links) for causal effect estimation, we use the causal graphs discovered
414 using TCDF in this study. Assuming the relationships to be linear with no interactions, the
415 dependence of Y on X and confounders C can be expressed mathematically as:

$$416 \quad E[Y|do(X = x), C = c] = \alpha X + \beta C + \gamma \quad (4)$$

417 Here, we use a doubly robust nonparametric estimator based on the theory of influence functions
418 called generalized augmented inverse probability weighted (gAIPW) estimator.^{28,29} We present
419 ACE along with its 95% confidence interval estimated using 500 bootstrap samples.

420 **Data availability**

421 The surface ozone data are available from GAW-WDCRG (<http://ebas.nilu.no>). The data for
422 solar flux, MEI and AAO index are taken from
423 (www.esrl.noaa.gov/psd/data/climateindices/list/). QBO data is taken from (<https://www.geo.fu-berlin.de/met/ag/strat/produkte/qbo/qbo.dat>). Heat flux is calculated using ERA-Interim
424 meteorological reanalyses from European Centre for Medium-Range Weather Forecasts
425 (ECMWF) and the AOD data is from Modern-Era Retrospective analysis for Research and
426 Applications, Version 2 (MERRA-2) reanalyses.

428 **Acknowledgement**

429 We thank the Head CORAL, the Director of Indian Institute of Technology Kharagpur (IIT
430 KGP), and the Ministry of Education (MoE) for facilitating the study. PK acknowledges the
431 support from MoE and IIT KGP. We thank the data managers and the scientists who worked
432 hard for making available ground-based surface ozone and all other data for this study.

433 **Author contributions**

434 PK conceived the idea, designed the research and performed the data analyses. PK wrote the first
435 draft, which was subsequently revised with inputs from JK and AM. JK supervised the research
436 at IIT KGP.

437 **Competing interests**

438 The authors declare no conflict of interest.

439 **Supplementary information**

440 Supplementary table and figures are available in the attached supplementary document.

441 **References**

- 442 1. Kuttippurath, J., Kumar, P., Nair, P. J. & Chakraborty, A. Accuracy of satellite total column
443 ozone measurements in polar vortex conditions: Comparison with ground-based observations in
444 1979–2013. *Remote Sensing of Environment* **209**, 648–659 (2018).
- 445 2. Solomon, S. *et al.* Emergence of healing in the Antarctic ozone layer. *Science* **353**, 269–274
446 (2016).
- 447 3. Kuttippurath, J., Kumar, P., Nair, P. J. & Pandey, P. C. Emergence of ozone recovery
448 evidenced by reduction in the occurrence of Antarctic ozone loss saturation. *npj Climate and*
449 *Atmospheric Science* **1**, 42 (2018).

- 450 4. Banerjee, A., Fyfe, J. C., Polvani, L. M., Waugh, D. & Chang, K.-L. A pause in Southern
451 Hemisphere circulation trends due to the Montreal Protocol. *Nature* 2020 579:7800 **579**, 544–
452 548 (2020).
- 453 5. Yang, J., Liu, J., Han, S., Yao, Q. & Cai, Z. Study of the meteorological influence on ozone in
454 urban areas and their use in assessing ozone trends in all seasons from 2009 to 2015 in Tianjin,
455 China. *Meteorology and Atmospheric Physics* **131**, 1661–1675 (2019).
- 456 6. Li, K. *et al.* Anthropogenic drivers of 2013–2017 trends in summer surface ozone in China.
457 *Proceedings of the National Academy of Sciences of the United States of America* **116**, 422–427
458 (2019).
- 459 7. Kretschmer, M., Coumou, D., Donges, J. F. & Runge, J. Using causal effect networks to
460 analyze different arctic drivers of midlatitude winter circulation. *Journal of Climate* **29**, 4069–
461 4081 (2016).
- 462 8. Donges, J. F. *et al.* Unified functional network and nonlinear time series analysis for complex
463 systems science: The pyunicorn package. *Chaos: An Interdisciplinary Journal of Nonlinear*
464 *Science* **25**, 113101 (2015).
- 465 9. Runge, J., Nowack, P., Kretschmer, M., Flaxman, S. & Sejdinovic, D. Detecting and
466 quantifying causal associations in large nonlinear time series datasets. *Science Advances* **5**,
467 eaau4996 (2019).
- 468 10. Cohen, L., Dean, S. & Renwick, J. Synoptic Weather Types for the Ross Sea Region,
469 Antarctica. *Journal of Climate* **26**, 636–649 (2013).
- 470 11. Neff, P. D. & Bertler, N. A. N. Trajectory modeling of modern dust transport to the Southern
471 Ocean and Antarctica. *Journal of Geophysical Research: Atmospheres* **120**, 9303–9322 (2015).
- 472 12. Masclin, S., Frey, M. M., Rogge, W. F. & Bales, R. C. Atmospheric nitric oxide and ozone at
473 the WAIS Divide deep coring site: a discussion of local sources and transport in West Antarctica.
474 *Atmospheric Chemistry and Physics* **13**, 8857–8877 (2013).

- 475 13. Legrand, M. *et al.* Inter-annual variability of surface ozone at coastal (Dumont d'Urville,
476 2004–2014) and inland (Concordia, 2007–2014) sites in East Antarctica. *Atmospheric Chemistry
477 and Physics* **16**, 8053–8069 (2016).
- 478 14. Kuttippurath, J. & Nair, P. J. The signs of Antarctic ozone hole recovery. *Scientific Reports*
479 *2017 7:1* **7**, 585 (2017).
- 480 15. Kuttippurath, J., Bodeker, G. E., Roscoe, H. K. & Nair, P. J. A cautionary note on the use of
481 EESC-based regression analysis for ozone trend studies. *Geophysical Research Letters* **42**, 1–7
482 (2014).
- 483 16. De Laat, A. T. J., Van Der A, R. J. & Van Weele, M. Tracing the second stage of ozone
484 recovery in the Antarctic ozone-hole with a "big data" approach to multivariate regressions.
485 *Atmospheric Chemistry and Physics* **15**, 79–97 (2015).
- 486 17. Balashov, N. V., Thompson, A. M., Piketh, S. J. & Langerman, K. E. Surface ozone
487 variability and trends over the South African Highveld from 1990 to 2007. *Journal of
488 Geophysical Research: Atmospheres* **119**, 4323–4342 (2014).
- 489 18. Kuchar, A. *et al.* On the aliasing of the solar cycle in the lower stratospheric tropical
490 temperature. *Journal of Geophysical Research: Atmospheres* **122**, 9076–9093 (2017).
- 491 19. Mansfield, L. A. *et al.* Predicting global patterns of long-term climate change from short-
492 term simulations using machine learning. *npj Climate and Atmospheric Science* **3**, (2020).
- 493 20. Kretschmer, M., Runge, J. & Coumou, D. Early prediction of extreme stratospheric polar
494 vortex states based on causal precursors. *Geophysical Research Letters* **44**, 8592–8600 (2017).
- 495 21. Nowack, P., Runge, J., Eyring, V. & Haigh, J. D. Causal networks for climate model
496 evaluation and constrained projections. *Nature Communications* **11**, 1–11 (2020).
- 497 22. Runge, J. *et al.* Inferring causation from time series in Earth system sciences. *Nature*
498 *Communications* **10**, 1–13 (2019).
- 499 23. Krich, C. *et al.* Estimating causal networks in biosphere-atmosphere interaction with the
500 PCMCI approach. *Biogeosciences* **17**, 1033–1061 (2020).

- 501 24. Runge, J. *et al.* Identifying causal gateways and mediators in complex spatio-temporal
502 systems. *Nature Communications* **6**, 1–10 (2015).
- 503 25. Bussmann, B., Nys, J. & Latré, S. Neural Additive Vector Autoregression Models for Causal
504 Discovery in Time Series Data. (2020).
- 505 26. Tank, A., Covert, I., Foti, N., Shojaie, A. & Fox, E. B. Neural Granger Causality. *IEEE*
506 *Transactions on Pattern Analysis and Machine Intelligence* 1–14 (2021).
507 doi:10.1109/TPAMI.2021.3065601
- 508 27. Nauta, M., Bucur, D. & Seifert, C. Causal Discovery with Attention-Based Convolutional
509 Neural Networks. *Machine Learning and Knowledge Extraction* **1**, 312–340 (2019).
- 510 28. Bhattacharya, R., Nabi, R. & Shpitser, I. Semiparametric Inference For Causal Effects In
511 Graphical Models With Hidden Variables. *arXiv* 1–73 (2020).
- 512 29. Henckel, L., Perković, E. & Maathuis, M. H. Graphical criteria for efficient total effect
513 estimation via adjustment in causal linear models. *arXiv* 1–58 (2019).
- 514 30. Lee, H.-M., Henze, D. K., Alexander, B. & Murray, L. T. Investigating the sensitivity of
515 surface-level nitrate seasonality in Antarctica to primary sources using a global model.
516 *Atmospheric Environment* **89**, 757–767 (2014).
- 517 31. Cristofanelli, P. *et al.* Analysis of multi-year near-surface ozone observations at the
518 WMO/GAW ‘Concordia’ station (75°06’S, 123°20’E, 3280 m a.s.l. – Antarctica). *Atmospheric*
519 *Environment* **177**, 54–63 (2018).
- 520 32. Gaudel, A. *et al.* Tropospheric Ozone Assessment Report: Present-day distribution and trends
521 of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation
522 Tropospheric Ozone Assessment Report: Present-day distribution and trends of tropospheric.
523 *Elem Sci Anth* (2018). doi:10.1525/elementa.291
- 524 33. Monks, P. S. *et al.* Tropospheric ozone and its precursors from the urban to the global scale
525 from air quality to short-lived climate forcer. *Atmospheric Chemistry and Physics* **15**, 8889–8973
526 (2015).

- 527 34. Cooper, O. R. *et al.* Multi-decadal surface ozone trends at globally distributed remote
528 locations. *Elementa: Science of the Anthropocene* **8**, 23 (2020).
- 529 35. Cooper, O. R. *et al.* Increasing springtime ozone mixing ratios in the free troposphere over
530 western North America. *Nature* **463**, (2010).
- 531 36. Stohl, A. & Sodemann, H. Characteristics of atmospheric transport into the Antarctic
532 troposphere. *Journal of Geophysical Research Atmospheres* **115**, 1–16 (2010).
- 533 37. Wang, S. *et al.* How Do Weakening of the Stratospheric Polar Vortex in the Southern
534 Hemisphere Affect Regional Antarctic Sea Ice Extent? *Geophysical Research Letters* **48**, 1–10
535 (2021).
- 536 38. Hirano, S., Kohma, M. & Sato, K. A three-dimensional analysis on the role of atmospheric
537 waves in the climatology and interannual variability of stratospheric final warming in the
538 Southern Hemisphere. *Journal of Geophysical Research* **121**, 8429–8443 (2016).
- 539 39. Calvo, N., Polvani, L. M. & Solomon, S. On the surface impact of Arctic stratospheric ozone
540 extremes. *Environmental Research Letters* **10**, 094003 (2015).
- 541 40. Screen, J. A., Bracegirdle, T. J. & Simmonds, I. Polar Climate Change as Manifest in
542 Atmospheric Circulation. *Current Climate Change Reports* **4**, 383–395 (2018).
- 543 41. Gonzalez, S., Vasallo, F., Recio-Blitz, C., Gujjarro, J. A. & Riesco, J. Atmospheric Patterns
544 over the Antarctic Peninsula. *Journal of Climate* **31**, 3597–3608 (2018).
- 545 42. Wespes, C., Hurtmans, D., Clerbaux, C. & Coheur, P.-F. O₃ variability in the troposphere as
546 observed by IASI over 2008-2016: Contribution of atmospheric chemistry and dynamics.
547 *Journal of Geophysical Research: Atmospheres* **122**, 2429–2451 (2017).
- 548 43. Kumar, P. *et al.* The Increasing Surface Ozone and Tropospheric Ozone in Antarctica and
549 Their Possible Drivers. *Environmental Science & Technology* (2021).
550 doi:10.1021/acs.est.0c08491

- 551 44. England, M. R., Shaw, T. A. & Polvani, L. M. Troposphere-stratosphere dynamical coupling
552 in the southern high latitudes and its linkage to the Amundsen Sea. *Journal of Geophysical*
553 *Research* **121**, 3776–3789 (2016).
- 554 45. Boljka, L. & Birner, T. Tropopause-level planetary wave source and its role in two-way
555 troposphere–stratosphere coupling. *Weather and Climate Dynamics* **1**, 555–575 (2020).
- 556 46. Pohl, B., Fauchereau, N., Reason, C. J. C. & Rouault, M. Relationships between the
557 Antarctic oscillation, the Madden-Julian oscillation, and ENSO, and consequences for rainfall
558 analysis. *Journal of Climate* **23**, 238–254 (2010).
- 559 47. Carvalho, L. M. V., Jones, C. & Ambrizzi, T. Opposite phases of the Antarctic oscillation
560 and relationships with intraseasonal to interannual activity in the tropics during the austral
561 summer. *Journal of Climate* **18**, 702–718 (2005).
- 562 48. Taguchi, M. Observed connection of the stratospheric quasi-biennial oscillation with El
563 Niño-Southern Oscillation in radiosonde data. *Journal of Geophysical Research Atmospheres*
564 **115**, 1–12 (2010).
- 565 49. Lu, X. *et al.* Surface and tropospheric ozone trends in the Southern Hemisphere since 1990:
566 possible linkages to poleward expansion of the Hadley circulation. *Science Bulletin* **64**, 400–409
567 (2019).
- 568 50. Löwe, S., Madras, D., Zemel, R. & Welling, M. Amortized Causal Discovery: Learning to
569 Infer Causal Graphs from Time-Series Data. (2020).

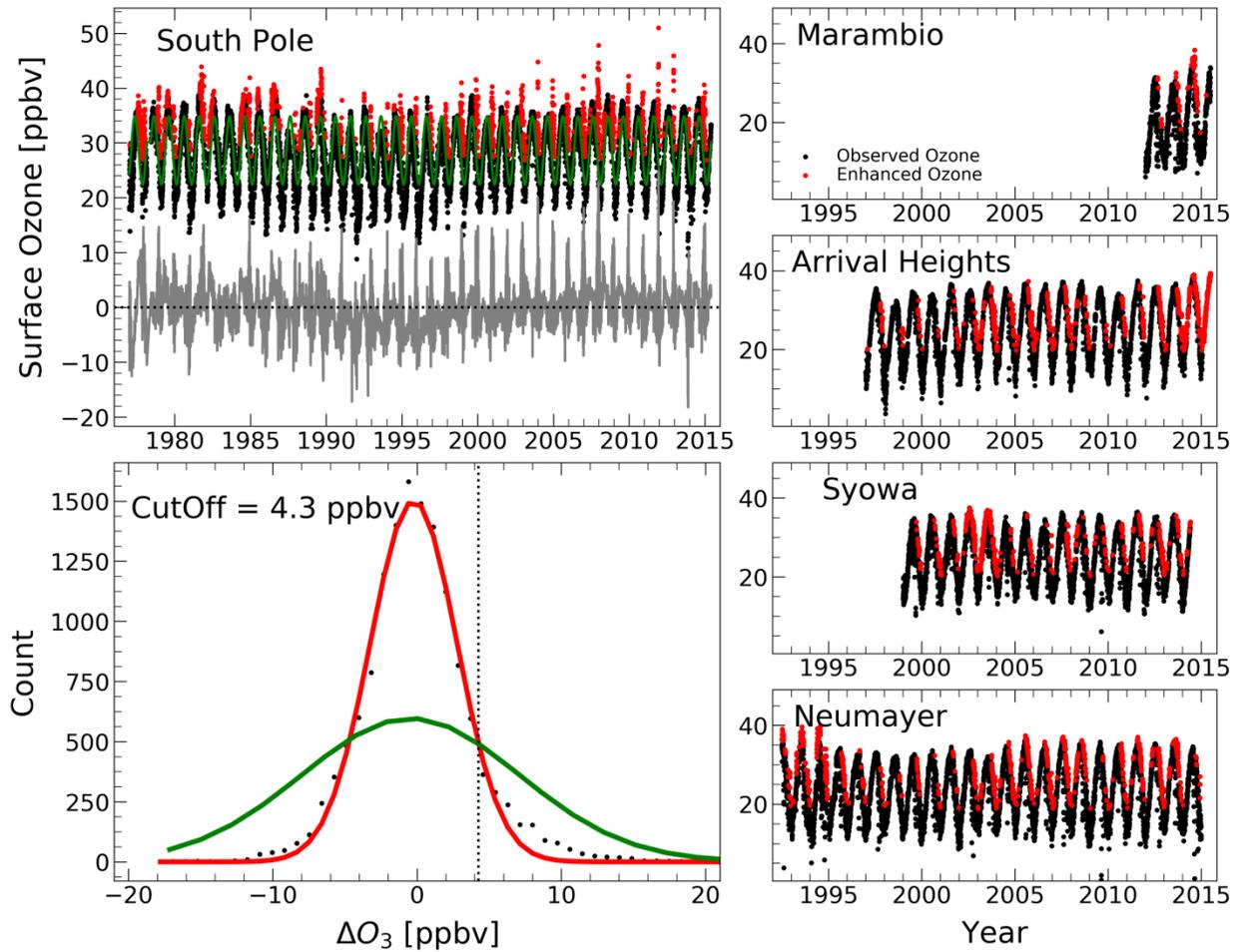
570

571 **Table 1:** Causal Effects (the proportion of changes in surface ozone with the change of 1
572 standard deviation (sd) of the geophysical driver under investigation) of various causal relations
573 discovered by TCDF estimated using generalized augmented inverse probability weighting
574 (gAIPW). Here average causal effect (ACE) is shown along with its 95% confidence interval
575 estimated by drawing 500 bootstrap samples.

	Arrival Heights	Neumayer	South Pole	Syowa
PV	0.93 [0.84, 1.04]	0.58 [0.53, 0.66]	0.77 [0.70, 0.86]	0.90 [0.82, 1.01]
AAO	-0.18 [-0.35, 0.01]	-0.13 [-0.29, 0.04]	-0.14 [-0.29, -0.01]	-0.32 [-0.54, -0.09]
HF	0.14 [-0.06, 0.36]	0.17 [0.02, 0.35]	0.12 [-0.03, 0.28]	0.24 [0.01, 0.52]
AOD	-2.51 [-4.26, -0.58]	-2.85 [-4.54, -1.26]	2.50 [0.71, 4.42]	-3.75 [-5.97, -1.39]

576

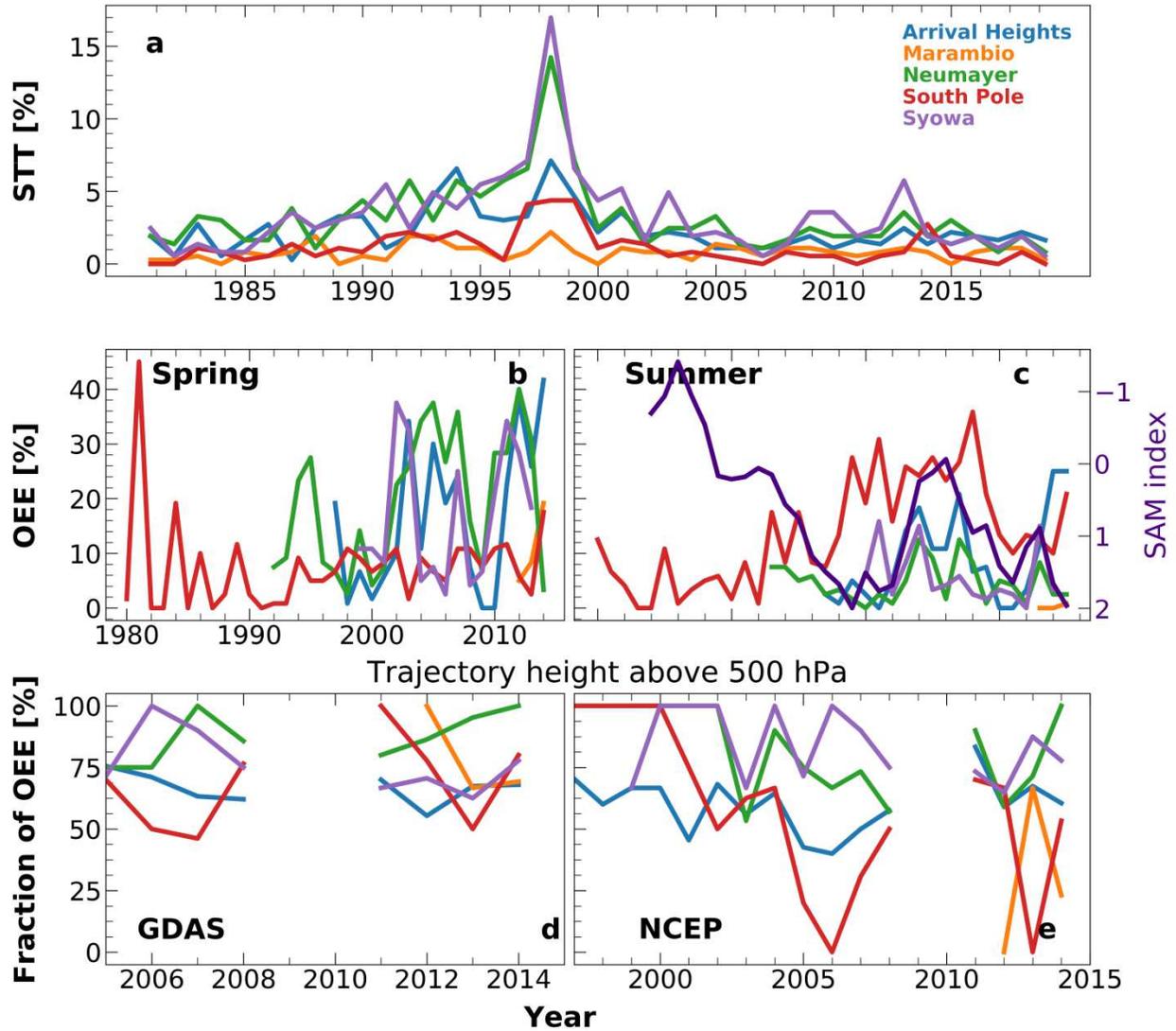
577



578

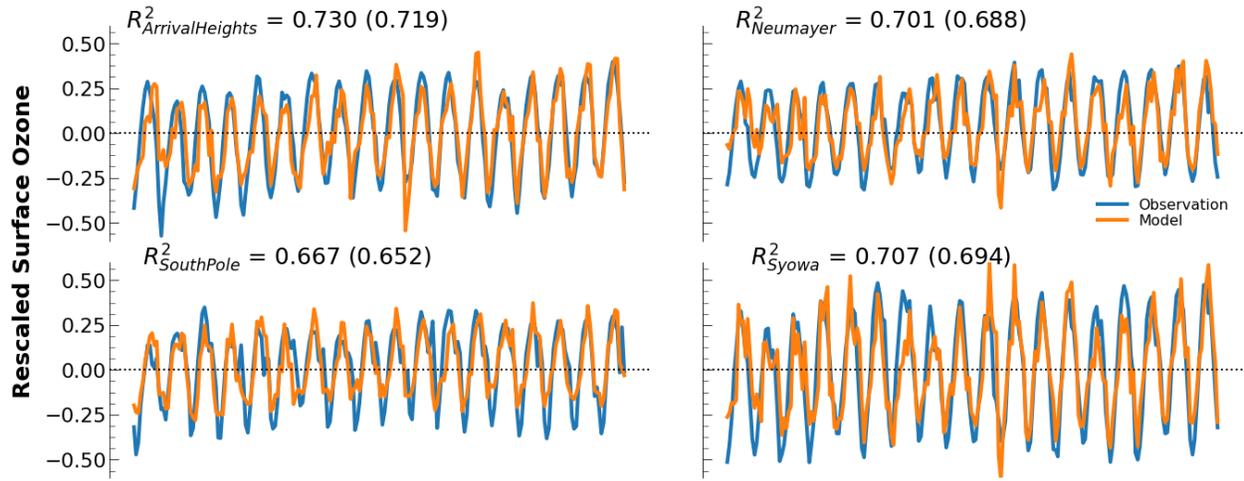
579 **Figure 1:** Daily time-series of surface ozone at various stations in Antarctica. Here, points in
580 magenta color show the identified enhanced ozone events (OEEs) using the methodology
581 described in the text.

582



583

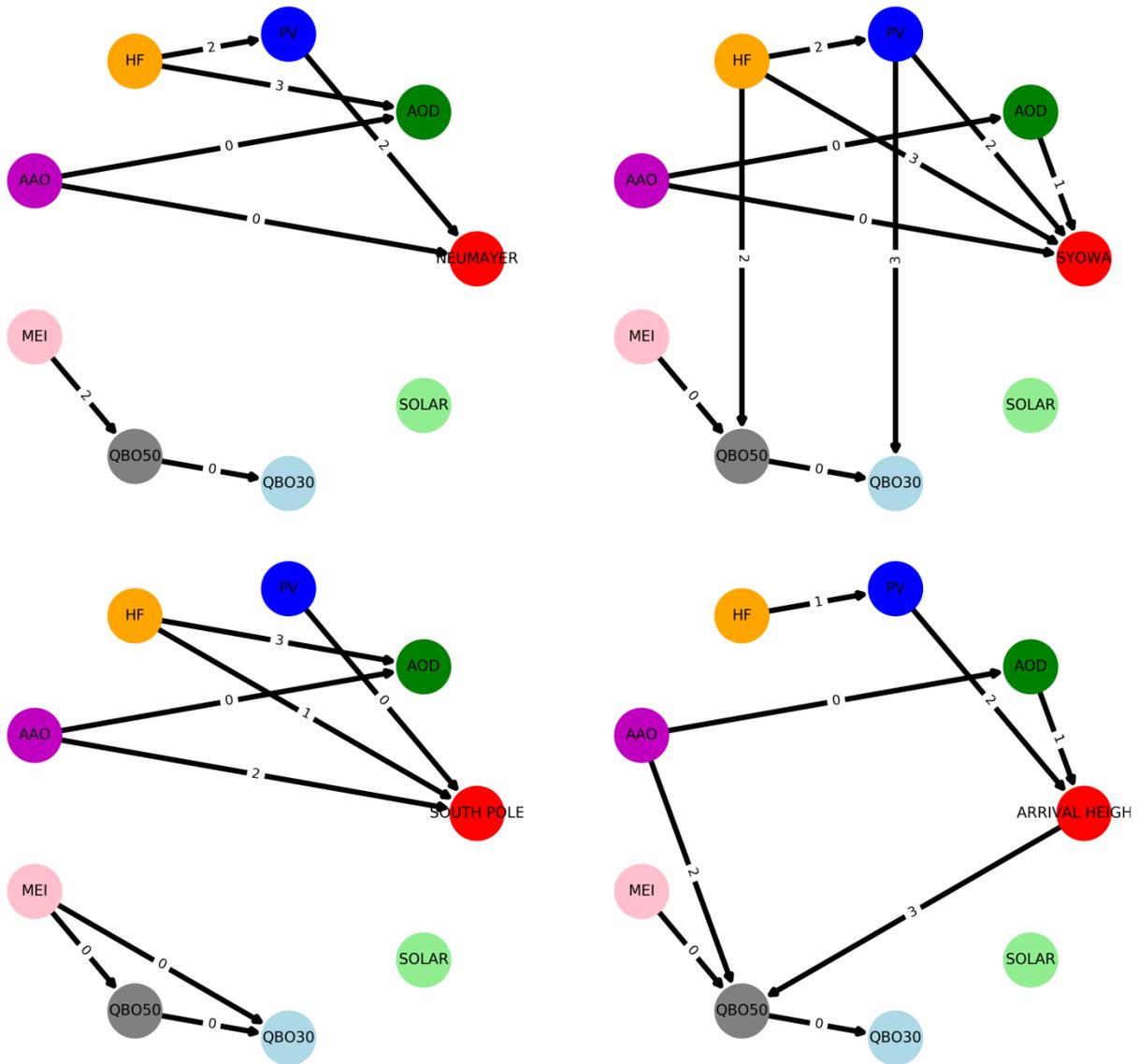
584 **Figure 2:** Inter-annual variation in the frequency of OEEs and associated back-trajectories
 585 altitude. a) Timeseries for OEEs during Spring. b) OEEs occurring during Summer. Here, the
 586 fraction of trajectories coming from UTLS (crossing 500 hPa) simulated using c) GDAS and d)
 587 NCEP meteorological reanalyses corresponding to OEEs is also shown.



588

589 **Figure 3:** MLR fit for Surface Ozone at different stations in Antarctica. Here, the goodness of fit
 590 is represented with R^2 and adjusted R^2 (R^2_{adj}).

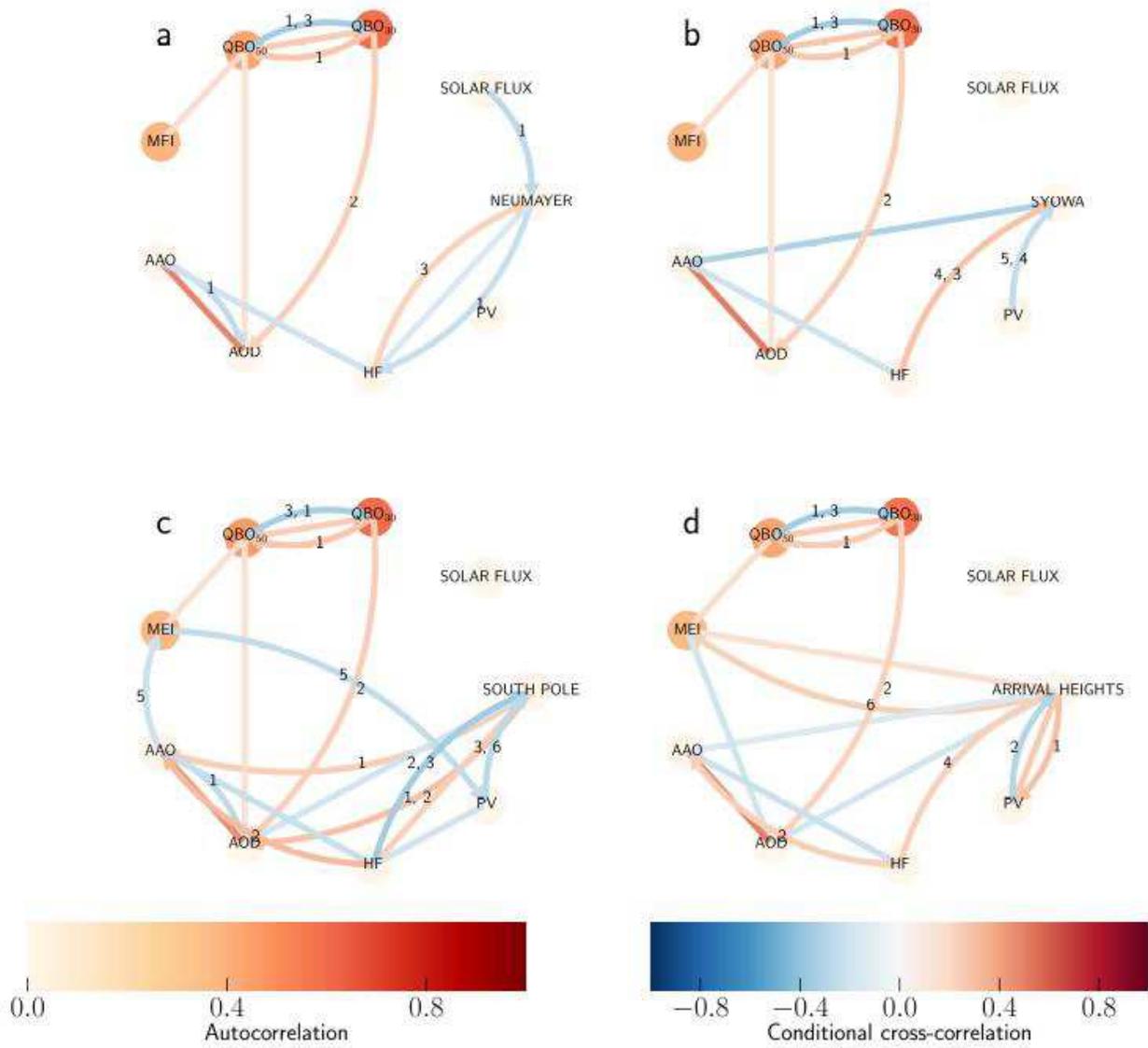
591



592

593 **Figure 4:** TCDF Causal Graph for Surface Ozone at all four stations (Neumayer, Syowa, Arrival
 594 Heights, and the South Pole) in Antarctica. Here, numbers in the middle of the detected links
 595 represent the optimal lag between cause and effect.

596



597

598 **Figure 5:** Causal graph for surface ozone at all four stations (Neumayer, Syowa, Arrival Heights
 599 and South Pole) considered in this study generated using PCMCI at 5% significance level. Here,
 600 the color of nodes shows the autocorrelation, whereas the color of the detected links represents
 601 the conditional cross-correlation between concerned nodes. The numbers in the middle of the
 602 detected links represent the detected lags between cause and effect.

603

604

605

Figures

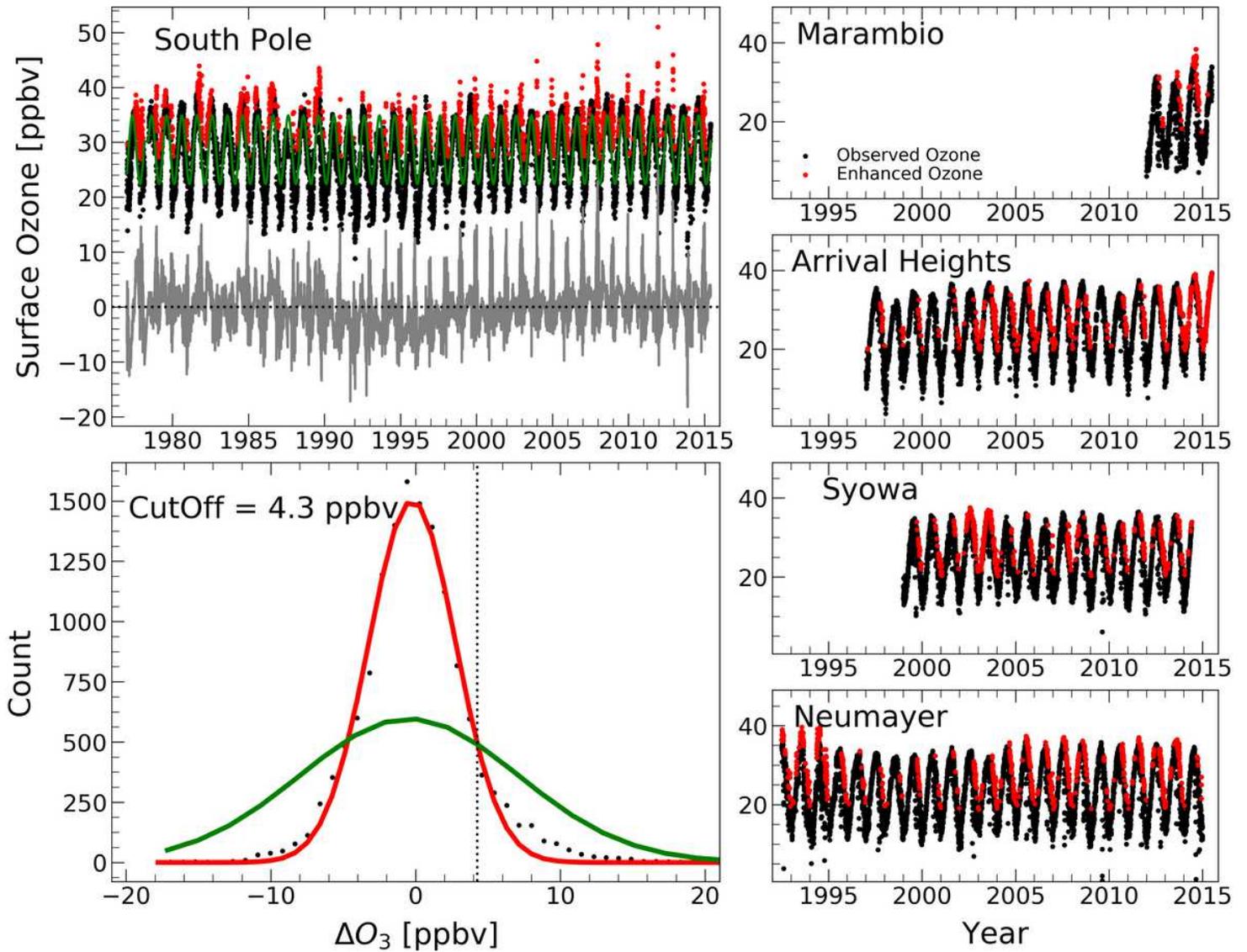


Figure 1

Daily time-series of surface ozone at various stations in Antarctica. Here, points in magenta color show the identified enhanced ozone events (OEEs) using the methodology described in the text.

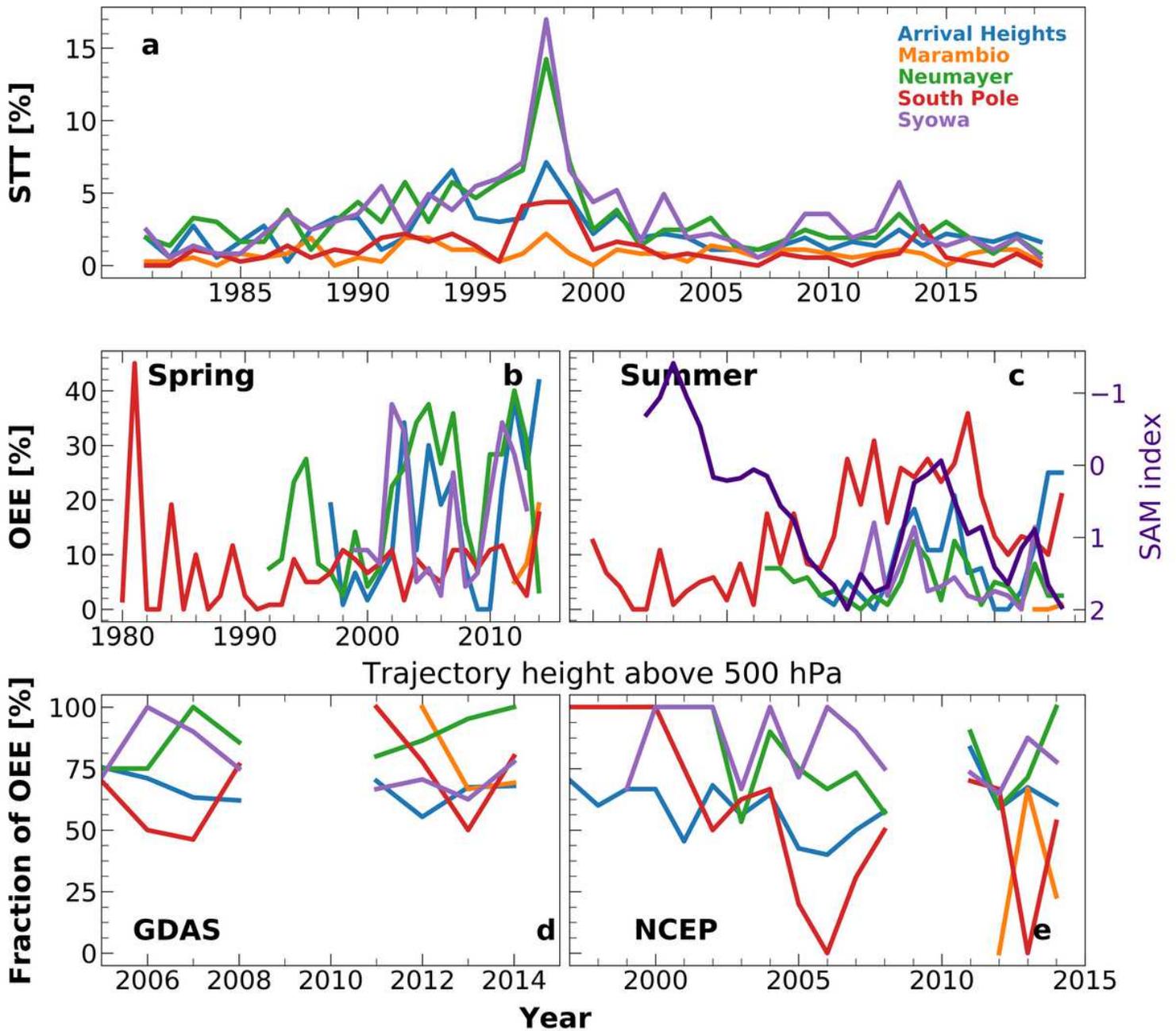


Figure 2

Inter-annual variation in the frequency of OEEs and associated back-trajectories altitude. a) Timeseries for OEEs during Spring. b) OEEs occurring during Summer. Here, the fraction of trajectories coming from UTLS (crossing 500 hPa) simulated using c) GDAS and d) NCEP meteorological reanalyses corresponding to OEEs is also shown.

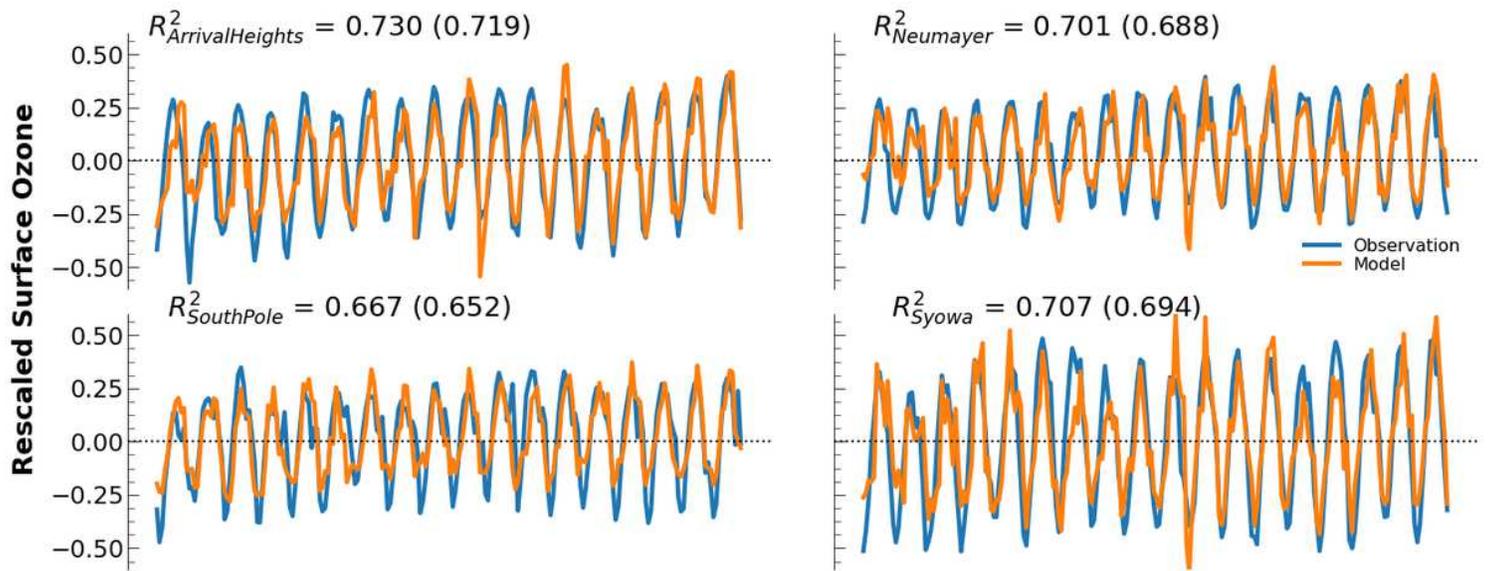


Figure 3

MLR fit for Surface Ozone at different stations in Antarctica. Here, the goodness of fit is represented with R^2 and adjusted R^2 ().

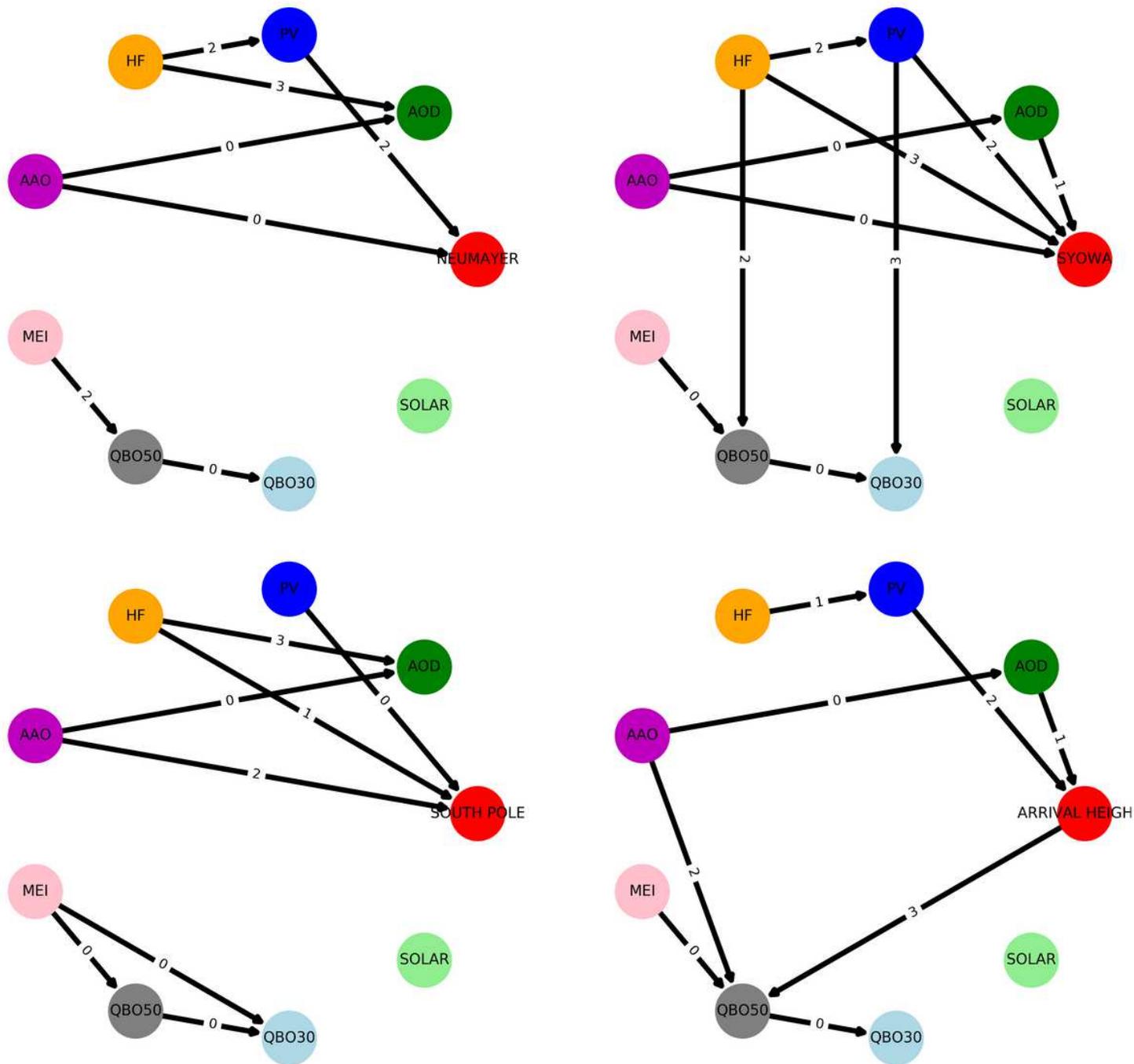


Figure 4

TCDF Causal Graph for Surface Ozone at all four stations (Neumayer, Syowa, Arrival Heights, and the South Pole) in Antarctica. Here, numbers in the middle of the detected links represent the optimal lag between cause and effect.

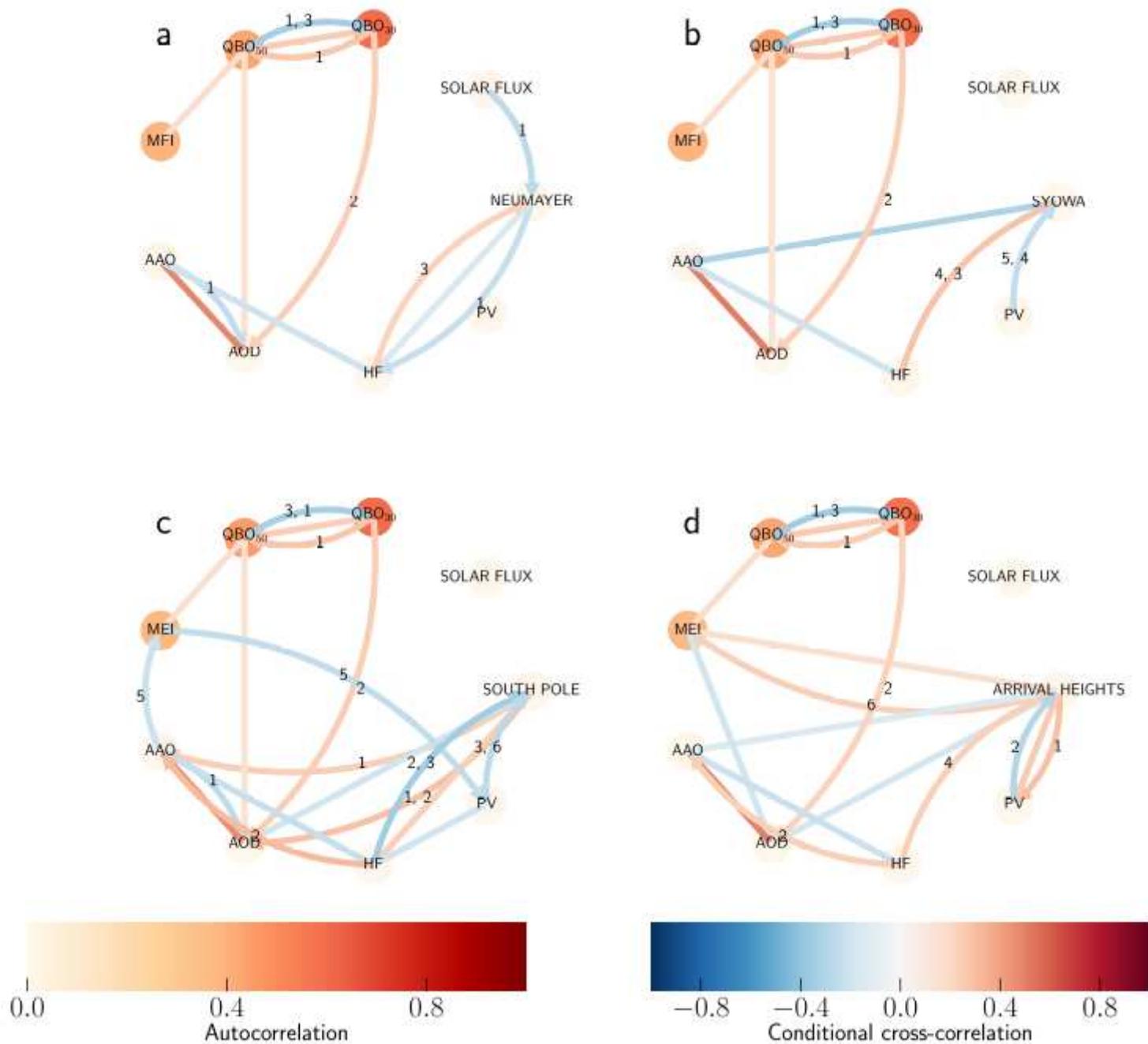


Figure 5

Causal graph for surface ozone at all four stations (Neumayer, Syowa, Arrival Heights and South Pole) considered in this study generated using PCMCI at 5% significance level. Here, the color of nodes shows the autocorrelation, whereas the color of the detected links represents the conditional cross-correlation between concerned nodes. The numbers in the middle of the detected links represent the detected lags between cause and effect.

Supplementary Files

This is a list of supplementary files associated with this preprint. [Click to download.](#)

- [causalsupp.docx](#)